Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **481,** 3661–3669 (2018)
Advance Access publication 2018 September 4

doi:10.1093/mnras/sty2393

Constraining sterile neutrino cosmologies with strong gravitational lensing observations at redshift $z \sim 0.2$

S. Vegetti,[1] G. Despali ,[1] M. R. Lovell[2,3] and W. Enzi[1]

[1] *Max-Planck Institute for Astrophysics, Karl-Schwarzschild Str. 1, D-85748 Garching, Germany*
[2] *Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*
[3] *Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

Accepted 2018 August 31. Received 2018 August 27; in original form 2017 December 22

ABSTRACT

We use the observed amount of subhaloes and line-of-sight dark matter haloes in a sample of 11 gravitational lens systems from the Sloan Lens ACS Survey to constrain the free-streaming properties of the dark matter particles. In particular, we combine the detection of a small-mass dark matter halo by Vegetti et al. with the non-detections by Vegetti et al. and compare the derived subhalo and halo mass functions with expectations from cold dark matter (CDM) and resonantly produced sterile neutrino models. We constrain the half-mode mass, i.e. the mass scale at which the linear matter power spectrum is reduced by 50 per cent relatively to the CDM model, to be $\log M_{\rm hm}[M_\odot] < 12.0$ (equivalent thermal relic mass $m_{\rm th} > 0.3$ keV) at the 2σ level. This excludes sterile neutrino models with neutrino masses $m_{\rm s} < 0.8$ keV at any value of L_6. Our constraints are weaker than currently provided by the number of Milky Way satellites, observations of the 3.5 keV X-ray line, and the Lyman α forest. However, they are more robust than the former as they are less affected by baryonic processes. Moreover, unlike the latter, they are not affected by assumptions on the thermal histories for the intergalactic medium. Gravitational lens systems with higher data quality and higher source and lens redshift are required to obtain tighter constraints.

Key words: gravitational lensing: strong – galaxies: haloes – galaxies: structure – dark matter.

1 INTRODUCTION

A significant number of astrophysical observations supports the widespread presence of dark matter and its dominant contribution to the matter content of the Universe. However, the nature of dark matter is to this day still an unsolved problem. In the standard ΛCDM cosmological model, dark matter is assumed to be cold, i.e. to have had negligible thermal velocity at early times, and possible candidates include the neutralino and the axions (e.g. Bottino, Fornengo & Scopel 2003; Ringwald 2016). While this model is very successful at reproducing large-scale observations, some discrepancies arise at galactic and sub-galactic scales (e.g. Moore 1994; Klypin et al. 1999; Kuzio de Naray, McGaugh & de Blok 2008; de Blok 2010; Walker & Peñarrubia 2011; Amorisco & Evans 2012; Boylan-Kolchin, Bullock & Kaplinghat 2012).

Concerning these issues, alternative models have been considered (e.g. Lovell et al. 2014; Vogelsberger et al. 2016; Iršič et al. 2017b; Robles et al. 2017), among which are the so-called warm dark matter models, where the corresponding particles have non-negligible thermal velocities. In particular, sterile neutrinos with masses of a few keV have been shown to be an interesting candidate for warm dark matter particles (e.g. Boyarsky, Iakubovskyi & Ruchayskiy 2012, and references therein). Moreover, the three sterile neutrino model, also known as the neutrino minimal standard model (νMSM), can also explain baryogenesis processes and neutrino oscillations (Boyarsky et al. 2012, and references therein). Recently, sterile neutrinos have gained renewed attention following the apparent detections of a 3.5 keV X-ray line in several galaxy clusters as well as M31, and at the centre of the Milky Way (Boyarsky et al. 2014; Bulbul et al. 2014; Boyarsky et al. 2015); for a review see, e.g. Iakubovskyi (2014). However, sterile neutrino decay as a source of the line remains contentious (e.g. Anderson, Churazov & Bregman 2015; Gu et al. 2015; Jeltema & Profumo 2016). Recent results from the Hitomi collaboration (Aharonian et al. 2017) have ruled out an atomic transition as the origin of the 3.5 keV line, but a dark matter interpretation is still possible due to the different sensitivity of the Hitomi instrument to narrow (atomic) and broad lines. The 11σ detection of the 3.5 keV signal by NuStar (Neronov, Malyshev & Eckert 2016) from a quiescent region of the sky and subsequent confirmation of this result with Chandra (Cappelluti et al. 2017) suggest that the line is not a statistical fluctuation nor an artefact of the *XMM–Newton*'s instrumentation. It is therefore important to find complimentary probes of these models as well as

⋆ E-mail: svegetti@mpa-garching.mpg.de

supporting future X-ray observatory missions such as XARM and ATHENA that will provide still stronger X-ray constraints.

Due to their appreciable thermal velocities at early times, warm dark matter particles free-stream out of density perturbations and are responsible therefore for a cut-off in the linear power spectrum and by extension in the formation of structures at small scales. The exact scale at which the suppression happens is strongly related to the dark matter particle production mechanism. For example, for 7 keV sterile neutrinos with a lepton asymmetry (defined in Section 3) much smaller or larger than $L_6 = 8$ this happens at the scale of progressively more massive dwarf galaxies (Lovell et al. 2016), while in a CDM model the suppression happens at much smaller scales. Quantifying the number of small mass structures is therefore an essential approach to distinguish among different dark matter models. For example, Kennedy et al. (2014) and Lovell et al. (2016) have used semi-analytic models to compare the abundance of luminous Milky Way satellites predicted by sterile neutrino dark matter models with the observed number of dwarf spheroidals, and derived constraints on the sterile neutrino mass and the lepton asymmetry L_6. However, since luminous satellites are a biased tracer of the underlying mass distribution, the interpretation of these results is complicated by assumptions made on baryonic processes. Moreover, the not so precisely known mass of the Milky Way is also a source of degeneracy in the number of expected satellites.

Recently, measurements of the Lyman α forest from high redshift QSO spectra have argued that all values of L_6 for sterile neutrino mass <10 keV are ruled out. However, these results are highly sensitive to assumptions made on the thermal histories for the intergalactic medium (Baur et al. 2017). It has been demonstrated by Garzilli, Boyarsky & Ruchayskiy (2017) that if one does not make any assumption about the (experimentally unknown) state of the intergalactic medium at $z > 5$, these bounds get considerably relaxed. As a result, the Lyman α limits are less stringent and in particular, the 7 keV sterile neutrino model, consistent with the 3.5 keV line detection, are within the allowed range (Baur et al. 2017).

Strong gravitational lensing, being sensitive to gravity, provides a more direct method to quantify the dark matter distribution at subgalactic scales and in a way is less affected by baryonic processes (however, see Despali & Vegetti 2017). Mao & Schneider (1998) first proposed the idea to constrain the amount of small-mass structure in gravitational lens galaxies via their influence on the relative fluxes of multiply-imaged quasars. Subsequently, Dalal & Kochanek (2002) were the first to use a relatively small sample of seven gravitationally lensed quasars to derive a statistical constraint on the projected mass fraction in substructure that is larger than, but consistent with, predictions from CDM numerical simulations. At the same time, analyses of CDM numerical simulations (Xu et al. 2009, 2015) have shown that the predicted substructure population is not sufficient to reproduce the level of flux-ratio anomalies currently observed, and that complex baryonic structures in the lens galaxies are a likely explanation (see also McKean et al. 2007; Hsueh et al. 2016, 2017a,b; Gilman et al. 2017). In the future, narrow-line observations for a large number of quadruply imaged quasars, which are expected to be discovered by large-scale surveys, will allow one to investigate these issues further and potentially provide new interesting constraints on the properties of dark matter (Nierenberg et al. 2017; Gilman et al. 2018).

In the meantime, Koopmans (2005) and Vegetti & Koopmans (2009) have introduced the *gravitational imaging* technique that uses magnified arcs and Einstein rings to detect and measure the mass of individual subhaloes (see also Vegetti, Czoske & Koopmans 2010a). Unlike other methods, this technique treats substructures not as analytical mass clumps but as pixellated linear potential corrections to the main lensing potential. As such, it does not require any prior assumption on the number of substructures nor their density profile and redshift. Moreover, it can easily distinguish a substructure from a smooth, but complex, mass distribution (Barnabè et al. 2009; Vegetti et al. 2014) and is therefore less prone to false detections. However, unlike studies of flux-ratio anomalies, it relies on increasingly high-resolution data for the detection of the smaller mass haloes. Using the gravitational imaging technique Vegetti et al. (2010b, 2012) have reported the detection of two small mass substructures, while Vegetti et al. (2014) have used a sample of 11 lenses to derive statistical constraints on the CDM substructure mass function, which are consistent with CDM predictions.

At the other end of the spectrum, several methods have been developed to detect mass substructures not individually, but via their collective gravitational effect (Fadely & Keeton 2012; Cyr-Racine et al. 2016; Hezaveh et al. 2016; Birrer, Amara & Refregier 2017; Daylan et al. 2017; Chatterjee & Koopmans 2018). Understanding the biases introduced by complex mass distributions of the main lensing galaxy, as highlighted by Hsueh et al. (2016, 2017a,b), Gilman et al. (2017), and Vegetti et al. (2014), will be critical for a practical application of these techniques.

Most if not all analyses so far have considered subhaloes within lens galaxies as the only source of perturbation to the lensed images. However, as demonstrated by Metcalf (2005) and Despali et al. (2018), the contribution from line-of-sight haloes can be significant and even dominant, depending on the lens and source redshifts. Moreover, both the number and the structure of line-of-sight haloes are less affected by feedback and accretion processes and can therefore be used to get tighter and cleaner constraints on the properties of dark matter.

In this paper we reanalyse the sample by Vegetti et al. (2014) to include the contribution of both subhaloes and line-of-sight haloes to the total number of detectable objects. In particular, we combine the non-detections by Vegetti et al. (2014) and the detection by Vegetti et al. (2010b) to derive statistical constraints on the halo and subhalo mass functions. As the key parameters that are set by the dark matter model are common to both mass functions, the inclusion of line-of-sight haloes represents an increase in the constraining power. This paper is organized as follows. The data is introduced in Section 2. In Section 3 we present the sterile neutrino dark matter model under consideration. In Sections 4 and 5 we provide all the definitions of mass for subhaloes and line-of-sight haloes as well as the mass function expressions adopted for both populations. In Sections 6 and 7 we introduce the complete likelihood function, the model parameters, and the corresponding priors. Finally, our results are presented and discussed in Sections 8 and 9, respectively.

2 DATA

The analysis presented in this paper is based on a sample of 11 gravitational lens systems taken from the SLACS survey (Bolton et al. 2006; Auger et al. 2010, and references therein). A summary of the systems with the lens and source redshift can be found in Table 1. Vegetti et al. (2010b) and Vegetti et al. (2014) have previously modelled this sample with the Bayesian grid-based adaptive code by Vegetti & Koopmans (2009). To summarize, Vegetti et al. (2010b) have reported the detection of a dark-matter-dominated perturber that was then interpreted as a substructure with a total mass of 3.5×10^9 $M_\odot$ (under the assumption of a Pseudo-Jaffe profile located on the plane of the host lens, hereafter PJ, and corresponding roughly to $M_{\rm vir}^{\rm NFW} \sim 10^{10} M_\odot$) in the lens system SDSSJ0946+1006.

Table 1. The list of the gravitational lens systems considered in this paper, with the lens and source redshifts.

Name (SDSS)	$z_{\rm lens}$	$z_{\rm source}$
J0252 + 0039	0.280	0.982
J0737 + 3216	0.322	0.581
J0946 + 1006	0.222	0.609
J0956 + 5100	0.240	0.470
J0959 + 4416	0.237	0.531
J1023 + 4230	0.191	0.696
J1205 + 4910	0.215	0.481
J1430 + 4105	0.285	0.575
J1627 − 0053	0.208	0.524
J2238 − 0754	0.137	0.713
J2300 + 0022	0.228	0.463

Later on, Vegetti et al. (2014) have shown that no substructure is required for all the other lenses in the sample. For each lens, they have also derived the so-called sensitivity function, that is the smallest detectable substructure mass at each relevant position on the image plane. Here, we make use of their results as an input to our analysis and refer to their papers for more details on the data and how these results were obtained. In particular, we relax the assumption that all perturbers are substructures and allow for the possibility of a contribution from line-of-sight haloes, for the statistical interpretation of both detections and non-detections.

3 DARK MATTER MODEL

The dark matter particle considered in this paper is the resonantly produced sterile neutrino (Shi & Fuller 1999; Dolgov & Hansen 2002; Asaka & Shaposhnikov 2005; Laine & Shaposhnikov 2008; Boyarsky, Ruchayskiy & Shaposhnikov 2009). In this section, we present a summary of the model and its application in astronomy. For an in-depth description of both, we refer the interested reader to Lovell et al. (2016).

The sterile neutrino is produced at high energies, around the gluon–hadron transition (∼10 MeV). It originates from the oscillation of active neutrinos, and the probability of this oscillation is enhanced in the presence of a lepton asymmetry, i.e. an excess of leptons over anti-leptons. It is parametrized using the L_6 parameter: $L_6 = 10^6(n_{\rm l} - \bar{n}_{\rm l})/s$, where $n_{\rm l}$ is the number density of leptons, $\bar{n}_{\rm l}$ the number density of anti-leptons, and s the entropy density. The ability to enhance the production of sterile neutrinos via lepton asymmetry reduces the required mixing angle between sterile neutrinos and active neutrinos to obtain the measured universal dark matter density, and thus evades bounds on sterile neutrino parameters derived from X-ray observations.

A second important consequence of lepton asymmetry-induced production is the effect on the sterile neutrino velocities. This is because the enhancement in the production rate due to the lepton asymmetry scales with the momentum of the neutrinos.

Small asymmetries ($L_6 < 6 - 25$, depending on the sterile neutrino mass) lead to excess production of low momentum sterile neutrinos and are therefore 'cooler' distributions than the no asymmetry case. At higher asymmetries, the enhancement in production is stronger and extends to high momentum sterile neutrinos. For the maximal lepton asymmetry ($L_6 > 100$) all momenta receive an equal boost compared to the zero asymmetry scenario and the resulting momentum distribution is therefore almost identical to that of no lepton asymmetry. Given that the mass of the sterile neutrino also plays a role in the momentum distribution, with less massive particles exhibiting higher velocities, the free-streaming scale, and thus the halo mass function, is specified by two parameters: sterile neutrino mass $m_{\rm s}$, and lepton asymmetry L_6.

The momentum distribution functions are calculated using the methods of Laine & Shaposhnikov (2008) and are used as inputs for Boltzmann solver codes to calculate the linear matter power spectrum, $P(k)$. In our case, we use matter power spectra that have been computed using a modified version of CAMB (Lewis, Challinor & Lasenby 2000), and examples of these are displayed in Fig. 4 of Lovell et al. (2016). These curves exhibit an array of cut-off slopes and positions. For this study we characterize each curve by the transfer function of the sterile neutrino linear matter power spectrum compared to its CDM counterpart, which is given by $T(k) = [P(k)_{\rm SN}/P(k)_{\rm CDM}]^{0.5}$, and specifically the wavenumber at which the transfer function has the value 0.5. This is known as the half-mode wavenumber, $k_{\rm hm}$. Its influence on the halo mass function is parametrized through a dependent parameter, the half-mode mass $M_{\rm hm}$, which takes the following form:

$$M_{\rm hm} = \frac{4\pi}{3}\bar{\rho}\left(\frac{\pi}{k_{\rm hm}}\right)^3, \quad (1)$$

where $\bar{\rho}$ is the present day mean matter density of the Universe. In practice, $M_{\rm hm}$ is a function of both $m_{\rm s}$ and L_6, thus placing constraints on the former will lead to limits on the sterile neutrino parameters. In some cases $M_{\rm hm}$ does not fully encapsulate the fine details of the matter power spectra, such as the shallower slopes of some models: we comment below on how a more accurate parametrization of these curves would likely change the results.

4 MASS DEFINITION

We assume the *true* mass of line-of-sight haloes and substructures, m, to be the virial mass of a Navarro, Frenk & White (1997, hereafter NFW) profile with a Duffy et al. (2008) concentration–mass relation. While this is a good description for the former, it is only an approximation for the latter. Despali et al. (2018) have shown that at fixed virial mass these have a larger concentration that is mildly dependent on the subhalo distance from the host centre. However, by comparing the different deflection angles, they have also shown that assuming a constant Duffy et al. (2008) mass–concentration relation plays a secondary role in terms of the lensing effect. In particular, this assumption leads to an error on the mass that is as low as 5 per cent for subhaloes with masses of 10^{5-6} $M_\odot$ and as large as 20 per cent for masses of 10^9 $M_\odot$; the error on the expected number of substructure is of the order of 10 per cent. We also assume that the concentration does not change with the dark matter model. As shown by Ludlow et al. (2016) the concentration of WDM haloes differs from the CDM case only at low masses, where the number of structures is strongly suppressed. Again, Despali et al. (2018) have shown that this assumption is of secondary importance. We refer to their paper for a more detailed discussion on the matter.

We assume that the *observed* mass $m^{\rm o}$ of perturbers, i.e. the mass that one would derive from the gravitational lens modelling of the data, and the lowest detectable mass are PJ total masses located on the plane of the host lens. We then use the mass–redshift relation derived by Despali et al. (2018) to statistically relate the *true* and *observed* masses to each other as described in Section 6.2. This approach follows from the fact that while the detections were made in a pixellated model independent way, they have been then characterized in terms of PJ substructures. The sensitivity function used in this paper and derived by Vegetti et al. (2014) has also been obtained under the assumption of PJ substructures.

5 DARK MATTER MASS FUNCTION

While Vegetti et al. (2014) have assumed a CDM model and have focused only on substructures, here, we allow for a more general dark matter model that includes the effect of particle free-streaming and the contribution from small-mass dark matter haloes located along the line of sight. Following Lovell et al. (2014) and Schneider et al. (2012), we parametrize the substructure and the halo mass function as follows:

$$n(m) = n(m)^{\rm CDM}\left(1+\frac{M_{\rm hm}}{m}\right)^{\beta}, \tag{2}$$

where the second factor expresses the effect of particle-free streaming. A slightly different parametrization was found to be a better fit for the subhalo mass function by Lovell et al. (2014):

$$n(m) = n(m)^{\rm CDM}\left(1+\frac{2.7\times M_{\rm hm}}{m}\right)^{\beta}. \tag{3}$$

In this paper, we make use of both, with the first one leading to a more significant number of WDM subhaloes and hence being more conservative. We assume the CDM mass function for substructures to be

$$n_{\rm sub}^{\rm CDM}(m) \propto m^{-\alpha}. \tag{4}$$

For the line-of-sight halo CDM mass function we adopt the expression by Sheth & Tormen (1999), with the best-fitting parameters optimized for the Planck cosmology calculated by Despali et al. (2016). As discussed in the previous section, for both populations the mass function is a function of the NFW virial mass for a fixed concentration–mass relation. Scatter in this relation is taken into account in a statistical sense as described in Section 6.2.

6 LIKELIHOOD FUNCTION

In this section, we derive an expression for the likelihood of detecting n perturbers (substructures plus line-of-sight haloes) with observed masses $\{m_1^{\rm ob},....,m_n^{\rm ob}\}$ at the projected positions $\{\mathbf{x}_1^{\rm ob},....,\mathbf{x}_n^{\rm ob}\}$, and no detection in all other mass and position ranges. Assuming a Poisson distribution we follow Marshall et al. (1983) and write the log-likelihood for a single lens galaxy as follows (see the Appendix for a derivation):

$$\begin{aligned}\log P\left(\{m_1^{\rm ob},....,m_n^{\rm ob}\},\{\mathbf{x}_1^{\rm ob},....,\mathbf{x}_n^{\rm ob}\}|\theta\right)\\ = -\int\left[\mu_s(m^{\rm o},\mathbf{x}^{\rm o})+\mu_l(m^{\rm o},\mathbf{x}^{\rm o})\right]dm^{\rm o}d\mathbf{x}^{\rm o}\\ +\sum_i^n\log\left[\mu_s\left(m_i^{\rm ob},\mathbf{x}_i^{\rm ob}\right)dm^{\rm o}d\mathbf{x}^{\rm o}+\mu_l\left(m_i^{\rm ob},\mathbf{x}_i^{\rm ob}\right)dm^{\rm o}d\mathbf{x}^{\rm o}\right].\end{aligned} \tag{5}$$

Here, θ is a vector containing the model-free parameters that define the (sub)structure mass function and is explicitly introduced in Section 7. $\mu_s\left(m^{\rm o},\mathbf{x}^{\rm o}\right)dm^{\rm o}d\mathbf{x}^{\rm o}$ and $\mu_l\left(m^{\rm o},\mathbf{x}^{\rm o}\right)dm^{\rm o}d\mathbf{x}^{\rm o}$ are the expected number of substructures and line-of-sight haloes, respectively, in the mass range $m^{\rm o}$, $m^{\rm o}+dm^{\rm o}$ and projected position range $\mathbf{x}^{\rm o}$, $\mathbf{x}^{\rm o}+d\mathbf{x}^{\rm o}$. A derivation of them is given in Section 6.2. The masses $m^{\rm o}$ in the above equations are intended as *observed* ones and are defined according to Section 4. These are integrated between the lowest detectable mass at each projected position, $M_{\rm low}^{\rm PJ}(\mathbf{x}^{\rm o})$, and $M_{\rm max}^{\rm PJ}=1.0\times10^{10}\,{\rm M}_\odot$. Both limits are intended as PJ total masses. It should be noted that here $\mathbf{x}^{\rm o}$ is the position on the plane of the main lens where the perturber can be detected, for substructures this corresponds with the projected position of the perturber, within a relatively small error. For line-of-sight haloes, due the multiple lens plane configuration, $\mathbf{x}^{\rm o}$ and $\mathbf{x}$ are related to each other via the recursive lens equation evaluated at $\mathbf{x}^{\rm o}$.

6.1 Sensitivity function

For each of the considered lenses, the substructure sensitivity function, that is the lowest detectable mass as a function of position on the image plane $M_{\rm low}^{\rm PJ}(\mathbf{x}^{\rm o})$, was derived by Vegetti et al. (2014). Briefly, this was calculated by identifying the smallest PJ total mass on the plane of the host lens responsible for a change of the Bayesian evidence by $\Delta\log E\leq-50$, relatively to a model with no substructure. Under the assumption of Gaussian noise, this corresponds to a 10σ detection. As a reference, the detections reported by Vegetti et al. (2010b) and Vegetti et al. (2012) were, respectively, at the 16σ and 12σ limit. As demonstrated by Vegetti et al. (2014) and Despali et al. (2018), considering a constant sensitivity across the image plane can have a significant impact on the expected number of detectable objects and therefore on the inferred mass function parameters. Thus, we have derived the sensitivity function for each pixel on the image plane with a signal-to-noise ratio (SNR) larger than three. In the following sections, we include its effect via $P(I=1|m^{\rm o},\mathbf{x}^{\rm o})$, which is expressed as

$$P(I=1|m^{\rm o},\mathbf{x}^{\rm o})=\begin{cases}1 & \text{if } m^{\rm o}\geq M_{\rm low}^{\rm PJ}(\mathbf{x}^{\rm o})\\ 0 & \text{otherwise}\end{cases}, \tag{6}$$

where I is a vector that is equal to one for detectable perturbers and zero otherwise. By definition, $P(I=1|m^{\rm o}=m_i^{\rm ob},\mathbf{x}^{\rm o}=\mathbf{x}_i^{\rm ob})=1$.

6.2 Expectation values

We now provide explicit expressions for the expectation values of substructures and line-of-sight haloes. Below, the integration limits are intended between $M_{\rm min}^{\rm NFW}=10^5M_\odot$ and $M_{\rm max}^{\rm NFW}=10^{11}M_\odot$ for the *true* NFW virial mass and within the SNR ≥3 region for the true projected position. For the redshift of line-of-sight haloes we integrate between the observer and the source, but exclude the region within the virial radius of the main lens, i.e. $z\in[z_{\rm lens}-10^{-4};z_{\rm lens}+10^{-4}]$. The substructure expectation value is given by

$$\begin{aligned}\mu_s(m^{\rm o},\mathbf{x}^{\rm o})=\mu_{0,s}\times\int P(I=1|m^{\rm o},\mathbf{x}^{\rm o})P(m^{\rm o}|m,z_{\rm lens})\\ \times P(m|\theta)P(\mathbf{x}^{\rm o}|\mathbf{x},z_{\rm lens})P(\mathbf{x})dmd\mathbf{x},\end{aligned} \tag{7}$$

where

$$P(m|\theta)=n(m|\theta)\left[\int n(m'|\theta)dm'\right]^{-1} \tag{8}$$

is such that $P(m|\theta)dm$ is the probability of finding one substructure in the mass range m, $m+dm$.

Introducing the projected dark matter mass of the primary lens $M_{\rm lens}$ within the region of interest and the corresponding projected dark matter mass fraction in substructure $f_{\rm sub}$ we can express $\mu_{0,s}$ as follows:

$$\mu_{0,s}=f_{\rm sub}M_{\rm lens}\left[\int m'P(m'|\theta){\rm d}m'\right]^{-1}. \tag{9}$$

In this paper, we define $f_{\rm sub}$ as the projected dark matter fraction in substructure with masses between $M_{\rm min}^{\rm NFW}$ and $M_{\rm max}^{\rm NFW}$ and within the considered region. In particular, $f_{\rm sub}$ is a mean value and is therefore the same for every galaxy in the sample. This assumption is not critical here, as we are considering a sample of lenses that is relatively homogeneous both in mass and redshift. This definition

differs from the one by Vegetti et al. (2014), as it uses a different definition of substructure mass, as well as a different parametrization of the substructure mass function.

Similarly, the expectation value for line-of-sight haloes is given by

$$\mu_l(m^{\rm o}, \mathbf{x}^{\rm o}) = \mu_{0,l} \times \int P(I=1|m^{\rm o}, \mathbf{x}^{\rm o}) P(m^{\rm o}|m, z) \times P(\mathbf{x}^{\rm o}|\mathbf{x}, z) P(m, z|\theta) P(\mathbf{x}) P(z)\, {\rm d}m{\rm d}z{\rm d}\mathbf{x}, \quad (10)$$

where $\mu_{0,l}$ is expressed as

$$\mu_{0,l} = \int n(m', z'|\theta) \frac{{\rm d}V(\mathbf{x}')}{{\rm d}z'} {\rm d}m'{\rm d}z'{\rm d}\mathbf{x}'. \quad (11)$$

$P(m, z|\theta){\rm d}m{\rm d}z$ is the probability of finding one line-of-sight halo in the mass range m, $m + {\rm d}m$, and in the redshift range z, $z + {\rm d}z$, and is related to the halo mass function as follows:

$$P(m, z|\theta){\rm d}m{\rm d}z = n(m, z|\theta)\frac{{\rm d}V}{{\rm d}z}{\rm d}m{\rm d}z \times \left[\int n(m', z'|\theta)\frac{{\rm d}V}{{\rm d}z'}{\rm d}m'{\rm d}z'\right]^{-1}. \quad (12)$$

As shown by Despali et al. (2018), the measurement error on the perturber positions is relatively small (i.e. within 2.5 times the PSF FWHM), hence for simplicity we assume $P(\mathbf{x}^{\rm o}|\mathbf{x}, z) = \delta(\mathbf{x} - g(\mathbf{x}^{\rm o}, z))$. For substructure $g(\mathbf{x}^{\rm o}, z) \equiv \mathbf{x}^{\rm o}$, for line-of-sight haloes $g(\mathbf{x}^{\rm o}, z)$ takes into account the effect of the recursive lens equation. Following the results by Xu et al. (2015) and Despali & Vegetti (2017) we assume a uniform probability for $P(\mathbf{x})$.

By comparing the lensing effect of PJ perturbers at the redshift of the main lens with those of NFW line-of-sight haloes and subhaloes, Despali et al. (2018) have derived a mass-redshift relation that allows one to map one population into the other; following their results and referring to the mass–redshift relation as $f(m, z)$ we define $P(m^{\rm o}|m, z)$ as follows:

$$P(m^{\rm o}|m, z) = \frac{1}{\sqrt{2\pi}m^{\rm o}\sigma(z)} \exp\left[-\frac{(\log m^{\rm o} - f(m, z))^2}{2\sigma^2(z)}\right]. \quad (13)$$

Essentially, for a line-of-sight halo of NFW virial mass m located at redshift z, $f(m, z)$ returns the PJ total mass situated on the plane of the main lens with the most similar gravitational lensing effect. Here, we do not use the mean relation reported by Despali et al. (2018), but we derive new values for the parameters (which depend on the primary lens model), for each of the lenses in our sample. The intrinsic scatter $\sigma(z)$ of the mass–redshift relation is also not the same as the one reported by Despali et al. (2018), but it is a sum in quadrature of the error on the observed mass and the uncertainty related to the scatter/different choice of the concentration–mass relation. The scatter $\sigma(z)$ does not account therefore for the measurement error on the main lens parameters. We have found this to be smaller than the scatter due to the perturber redshift degeneracy and mass density profile (Despali et al. 2018), once incorrect modelling of the main deflector (e.g. wrong parametrisation) has been ruled out or addressed.

7 PRIOR AND POSTERIOR PROBABILITIES

The target parameters of the model θ, include the slope of the substructure mass function α (equation 4), the mean projected mass fraction in substructures with virial masses between $M_{\rm min}^{\rm NFW}$ and $M_{\rm max}^{\rm NFW}$, $f_{\rm sub}$ (equation 9), the half-mode mass $M_{\rm hm}$, and the slope β (equations 2 and 4). Prior probabilities on these parameters are chosen as follows:

Table 2. Current and projected constraints. For the two mass function slopes, we report the mean value as well as the 68 and 95 per cent CL. For the substructure mass fraction, we report the 68 and 95 per cent upper limits, while for the half-mode mass we also report the corresponding lower limits. The top table shows the current constraints derived assuming a 10σ detection threshold for the calculation of the sensitivity function. The middle (bottom) table shows the projected results that one could obtain with a sample of lenses with a sensitivity that is 10 (100) times better than the current one, and a single detection with the same mass and position as in the original data.

Run	Parameter	mean	σ_{68}	σ_{95}
$M_{\rm low}^{\rm PJ}$				
	α	1.87	−0.20 | +0.18	−0.33 | +0.35
	β	−1.31	−0.09 | +0.09	−0.17 | +0.17
	$f_{\rm sub}$	–	<0.087	<0.16
	$\log M_{\rm hm}[{\rm M}_\odot]$	–	9.14 | 11.9	6.42 | 12.0
$M_{\rm low}^{\rm PJ}/10$				
	α	1.81	−0.17 | +0.14	−0.28 | +0.31
	β	−1.31	−0.09 | +0.09	−0.16 | +0.17
	$f_{\rm sub}$	–	<0.089	<0.17
	$\log M_{\rm hm}[{\rm M}_\odot]$	–	10.6 | 12.1	9.40 | 12.4
$M_{\rm low}^{\rm PJ}/100$				
	α	1.84	−0.17 | +0.17	−0.31 | +0.33
	β	1.31	−0.09 | +0.09	−0.16 | +0.17
	$f_{\rm sub}$	–	<0.074	<0.14
	$\log M_{\rm hm}[{\rm M}_\odot]$	–	10.6 | 12.1	9.57 | 12.4

(i) α is drawn from a normal prior density distribution centred on $\alpha = 1.9$ and with a standard deviation of 0.2. This is in agreement with dark matter-only and hydrodynamical numerical simulations (Despali & Vegetti 2017);

(ii) for the normalization $f_{\rm sub}$ of the substructure mass function, we assume a uniform prior density distribution proportional to $f^{-0.5}$ between 0 and 0.2;

(iii) we assume β to be normally distributed around $\beta = -1.3$ (Lovell et al. 2017) with a standard deviation of 0.1;

(iv) we adopt a logarithmic prior distribution between 10^6 and 2×10^{12} ${\rm M}_\odot$ for the half-mode mass.

We derive the posterior probability for the mass function parameters θ from the likelihood function (equation 5) assuming the detection of perturbers from one lens system to another to be independent events. We explore the posterior parameter space within the prior volume using MULTINEST (Feroz & Hobson 2008). Results are presented in the following section.

8 RESULTS

We summarize our constraints on the model parameters in Table 2. For the two slopes, α and β we report the mean values, the 68 and the 95 per cent confidence levels (CLs). For the substructure mass fraction, we only report the 68 and 95 per cent upper limits, the posterior distribution being skewed towards high values. Constraints on the half-mode mass are expressed in terms of 68 and 95 per cent lower and upper limits. The posterior probability distributions for $f_{\rm sub}$ and $M_{\rm hm}$ are shown in Fig. 1.

First of all, we notice that the 68 per cent upper limit on $f_{\rm sub}$ is larger than the value reported by Vegetti et al. (2014) for the same sample of lenses. This difference can be attributed to a different definition of the substructure mass and mass limits as well as a different shape of the substructure mass function, which introduces a degeneracy between $f_{\rm sub}$ and $M_{\rm hm}$.

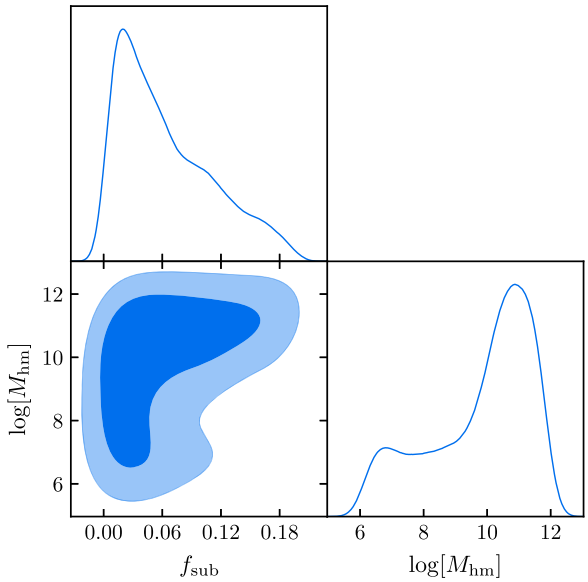

Figure 1. The posterior probability distribution for the projected substructure mass fraction and the half-mode mass derived by taking into account the contribution from both substructures and line-of-sight haloes. Contours correspond to the 1σ and 2σ levels.

At the 95 per cent CL, we constrain the half-mode mass to be $6.42 < \log M_{\rm hm}[{\rm M}_\odot] < 12.0$. These limits, although rather weak, are independent of the subhalo mass function parametrization, expressed by equations (2) and (3). Sterile neutrinos are a two-parameter dark matter model, where combinations of neutrino masses and lepton asymmetry in the early Universe determine the particle momentum distribution and its colder or warmer behaviour. In the left-hand panel of Fig. 2, we plot the theoretical half-mode mass for different values of the neutrino mass and lepton asymmetry. Our 2σ upper limit excludes sterile neutrino masses $m_{\rm s} < 0.8$ keV at any value of L_6. We have also derived a relationship between the mass of a thermal relic particle and the half-mode mass using the results of Viel et al. (2005), which leads to a lower limit of $m_{\rm th} > 0.3$ keV at the 2σ level.

In the right-hand panel of Fig. 2 we compare our constraints with those derived from the observed satellites in the Milky Way (Lovell et al. 2016), X-ray decay searches from M31 (Watson et al. 2012; Horiuchi et al. 2014), and Lyman α forest constraints. The latter measures the 1D matter power spectrum of Lyman α flux in QSO spectra. Comparing the limits from these studies with our results is complicated because their constraints are calculated using thermal relic matter power spectra, and a proper analysis requires simulations of structure formation to model the non-linear evolution of the power spectrum such as the flow from large scales to small scales. Another uncertainty is the thermal history around $z \sim 5$, where Iršič et al. (2017a) find their preferred power-law prior requires $m_{\rm th} > 5.3$ keV at the 95 per cent CL, whereas a freer prior on the thermal history relaxes the bound to >3.5 keV at 95 per cent CL.

We therefore take the following approach. We draw an exclusion region based on all sterile neutrino models that have a 1D power spectrum with less power at any point in the wavenumber range $1 < k < 10h$/Mpc than the 3.5 keV thermal relic, where $< 10\,h$/Mpc is the range of wavenumbers used in the analysis of Iršič et al. (2017a). When combined with the X-ray limit, this limit rules out all but a sliver of parameter space, which lies in the range $m_{\rm s} > 5$ keV, $L_6 \sim [8 - 10]$; the less conservative 5.3 keV limit instead rules out all $m_{\rm s} < 10$ keV. Finally, we note that our method may rule out models in which the power transfer from large scales to small scales is stronger than for the thermal relic, therefore dedicated simulations of these sterile neutrino models will be required to confirm or correct this simple model (see also Baur et al. 2016).

Our 95 per cent CL exclusion regions are significantly smaller than those derived from both the satellite counts and the Lyman α forest, and would potentially be weaker still if the shallower slopes of sterile neutrino power spectra were taken into account fully. However, they are more robust than those from the Milky Way satellite counts, as they are less affected by feedback processes.

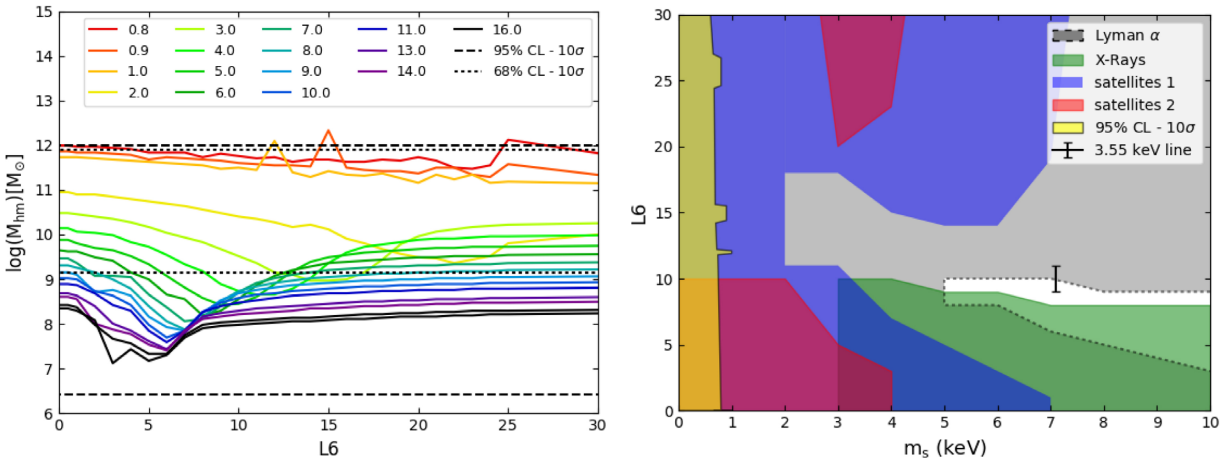

Figure 2. Current constraints – Left-hand panel: Half-mode mass versus lepton asymmetry for different values of the neutrino mass (coloured lines), 95 and 68 per cent upper and lower limits on the half-mode mass for the current sensitivity function (dashed and dotted black lines). Right-hand panel: 95 per cent exclusion region in the $L_6 - m_{\rm s}$ plane. The green shaded region is excluded from non-observations of X-ray decay from M31 (Watson, Li & Polley 2012; Horiuchi et al. 2014). The purple and blue regions are excluded by the observed number of Milky Way satellites for two different feedback models from Lovell et al. (2016). The grey shaded region is in strong tension with Lyman α flux observations as described in the text. The yellow shaded region is excluded by the number of observed and non-observed mass perturbers in the sample of gravitational lens systems considered in this paper. We mark the position of the sterile neutrino model that explains the 3.55 keV line with an error bar.

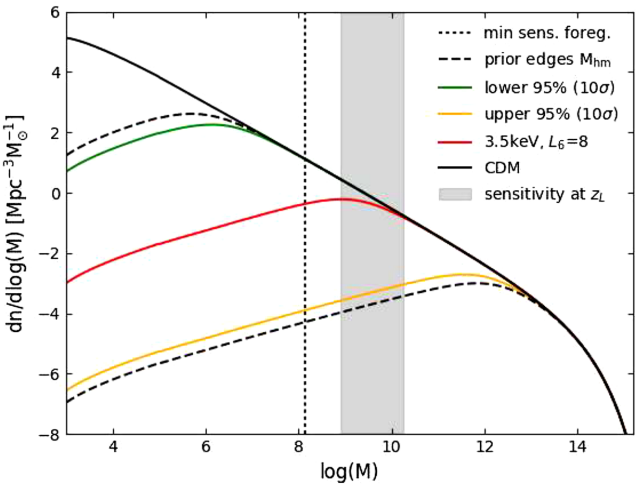

Figure 3. Line-of-sight differential mass functions according to the CDM (black line) model, the sterile neutrino model compatible with the detection of the 3.55 keV line (red), and the inferred 2σ upper (orange) and lower (green) limits. The grey area represents the region in substructure mass probed by the current detection threshold, and the vertical dotted line the lowest detectable mass for a foreground line-of-sight halo.

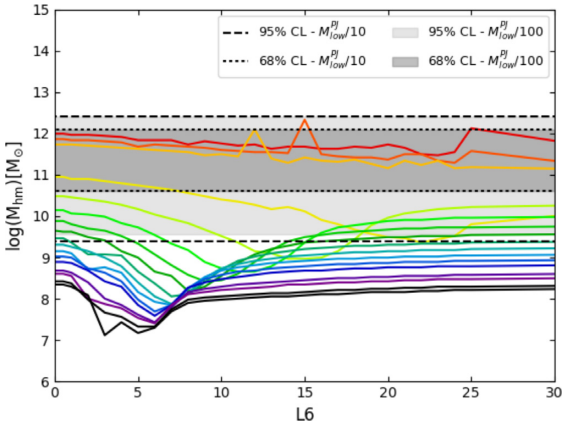

Figure 4. Projected constraints – Half-mode mass versus lepton asymmetry for different values of the neutrino mass (coloured lines), 95 and 68 per cent upper and lower limits on the half-mode mass for a sensitivity function artificially improved by one and two orders of magnitudes (dashed and dotted black lines, and grey bands). We have derived these constraints under the assumption of one detected perturber with the same mass and position of the one included in the original sample.

We note that the derived 95 per cent lower limit on $M_{\rm hm}$ is slightly larger than the lower limit imposed by our prior. However, this constraint is still prior dominated. In Fig. 3, we plot the differential line-of-sight mass function corresponding to the CDM model, the sterile neutrino model compatible with the 3.5 keV emission line, and the upper and lower limits derived in this paper. Within the mass limits imposed by the sensitivity of the data (expressed now as virial NFW mass), the CDM mass function (and therefore any model colder than $\log M_{\rm hm} = 6.0$) is virtually indistinguishable from the one corresponding to our lower limit. We can conclude therefore that our current results are not in tension with the prediction from CDM. Indeed the expected number of detectable line-of-sight haloes is 0.8 ± 0.9, in agreement with the single detection considered in this paper. In particular, we find that a set of data with a better sensitivity will be required to constrain models with $\log M_{\rm hm} < 6.0$. For example, as shown in Fig. 4 and Table 2, the same small sample of 11 lenses considered in this paper but with a sensitivity improved by one or two orders of magnitudes, would result in a posterior probability distribution for the half-mode mass which is significantly shifted towards larger values, and more stringent constraints on both limits of the half-mode mass. Under the assumption that only one perturber, with the same mass as the one reported by Vegetti et al. (2010b), is detected, such samples of lenses would allow us to rule out CDM at the 2σ level.

9 CONCLUSIONS

Sterile neutrinos with masses of a few keV have been shown to be a viable candidate for dark matter particles. Depending on the level of lepton asymmetry in the early Universe, these can behave as warmer or cooler models and introduce a cut-off in the primordial linear matter power spectrum. As a consequence, structure formation within sterile neutrino cosmologies is suppressed at a scale that is a function of the neutrino mass and the lepton asymmetry. In particular, 7 keV sterile neutrinos have been shown to be a possible explanation for the apparent 3.5 keV X-ray line detection in several galaxy clusters as well as M31 and at the centre of the Milky Way. Recently, Lovell et al. (2016) have used semi-analytical models with different sterile neutrino models to quantify the predicted number of luminous satellites in the Milky Way halo and concluded that the 7 keV sterile neutrinos are in good agreement with current observations. However, constraints on dark matter from the Milky Way satellites are limited by our knowledge of the Milky Way halo mass and the details of feedback processes.

Strong gravitational lensing allows one to gain new insight into the properties of dark matter by quantifying the number of substructure in gravitational lens galaxies and small-mass haloes along their line of sight, in a way that is less affected by galaxy formation models. In this paper, we have used the non-detections of mass perturbers obtained by Vegetti et al. (2014) in combination with the detection in the lens system SDSS J0946+1006 by Vegetti et al. (2010b) to derive constraints on the two-parameter model of sterile neutrinos. At the 2σ level we have excluded models with $m_{\rm s} < 0.8$ keV at any value of L_6.

These constraints are currently less stringent than those provided by other, more established methods. We have made explicit comparisons to the Milky Way satellite counts (Polisensky & Ricotti 2011; Kennedy et al. 2014; Lovell et al. 2016; Cherry & Horiuchi 2017), which almost always prefer a thermal relic mass >2 keV and, in the case of Lovell et al. (2016), set constraints on sterile neutrinos of mass <7 keV when applying their fiducial galaxy formation model; but their models with weaker reionization feedback can, however, evade the constraints. There are also uncertainties in the abundance and spatial distribution of the observed Milky Way satellites, in the halo mass and, the issue that this analysis is limited to the Milky Way and Andromeda systems.

Still stronger constraints are obtained from the analysis of the matter power spectrum of perturbations in the Lyman α forest (Schneider 2016). In particular, high redshift QSO spectra from XQ-100, HIRES, and MIKE have been argued to rule out all values of L_6 for sterile neutrino mass <10 keV; however, different thermal histories for the intergalactic medium return weaker limits (Baur et al. 2017). X-ray constraints instead set an upper limit on the sterile neutrino mass and a corresponding lower limit on L_6. The strongest limits, particularly for the 3.55 keV line, have arguably been obtained using deep *XMM–Newton* observations of the Draco dwarf spheroidal galaxy, but the long integration times required

for a detection, coupled to uncertainties in how best to conduct those observations, reduce the possibility that further studies can improve the situation in the near future (c.f. Jeltema & Profumo 2016; Ruchayskiy et al. 2016). Compared to all of the above methods, lensing studies benefit from the fact that they are sensitive to the total perturber mass and they are less affected by baryonic processes and should be therefore more robust. However, our current constraints are valid as long as the mass function parametrization adopted here is valid for any value of L_6 and m_s. We plan to test this hypothesis with numerical simulations of sterile neutrino models that span the range in L_6 inferred from the 3.5 keV line in a follow-up paper.

In agreement with Despali et al. (2018), we have found that including the contribution of line-of-sight haloes can provide a boost to the constraining power of gravitational lensing. However, due to the low redshift of the current sources and lenses, this effect is relatively small. In the near future, we will provide more stringent constraints on sterile neutrino models and the properties of dark matter in general, by using higher resolution data and samples of gravitational lens systems with a higher combination of source and lens redshift.

ACKNOWLEDGEMENTS

The authors are grateful to A. Boyarsky, O. Ruchayskiy, and S.D.M. White for useful comments and discussions. MRL is supported by a COFUND/Durham Junior Research Fellowship under EU grant 609412. MRL acknowledges support by a Grant of Excellence rom the Icelandic Research Fund (grant number 173929 − 051). SV has received funding from the European Research Council (ERC) under the Europen Union's Horizon 2020 research and innovation programme (grant agreement No 758853).

REFERENCES

Aharonian F. A. et al., 2017, ApJ, 837, L15
Amorisco N. C., Evans N. W., 2012, MNRAS, 419, 184
Anderson M. E., Churazov E., Bregman J. N., 2015, MNRAS, 452, 3905
Asaka T., Shaposhnikov M., 2005, Phys. Lett. B, 620, 17
Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Moustakas L. A., Burles S., 2010, ApJ, 724, 511
Barnabè M., Nipoti C., Koopmans L. V. E., Vegetti S., Ciotti L., 2009, MNRAS, 393, 1114
Baur J., Palanque-Delabrouille N., Yèche C., Magneville C., Viel M., 2016, J. Cosmol. Astropart. Phys., 8, 012
Baur J., Palanque-Delabrouille N., Yèche C., Boyarsky A., Ruchayskiy O., Armengaud É., Lesgourgues J., 2017, J. Cosmol. Astropart. Phys., 12, 013
Birrer S., Amara A., Refregier A., 2017, J. Cosmol. Astropart. Phys., 5, 037
Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, ApJ, 638, 703
Bottino A., Fornengo N., Scopel S., 2003, Phys. Rev. D, 67, 063519
Boyarsky A., Ruchayskiy O., Shaposhnikov M., 2009, Annu. Rev. Nucl. Part. Sci., 59, 191
Boyarsky A., Iakubovskyi D., Ruchayskiy O., 2012, Phys. Dark Universe, 1, 136
Boyarsky A., Ruchayskiy O., Iakubovskyi D., Franse J., 2014, Phys. Rev. Lett., 113, 251301
Boyarsky A., Franse J., Iakubovskyi D., Ruchayskiy O., 2015, Phys. Rev. Lett., 115, 161301
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, MNRAS, 422, 1203
Bulbul E., Markevitch M., Foster A., Smith R. K., Loewenstein M., Randall S. W., 2014, ApJ, 789, 13
Cappelluti N. et al., 2018, ApJ, 854, 179
Chatterjee S., Koopmans L. V. E., 2018, MNRAS, 474, 1762
Cherry J. F., Horiuchi S., 2017, Phys. Rev. D, 95, 083015
Cyr-Racine F.-Y., Moustakas L. A., Keeton C. R., Sigurdson K., Gilman D. A., 2016, Phys. Rev. D, 94, 043505
Dalal N., Kochanek C. S., 2002, ApJ, 572, 25
Daylan T., Cyr-Racine F.-Y., Diaz Rivero A., Dvorkin C., Finkbeiner D. P., 2018, ApJ, 854, 141
de Blok W. J. G., 2010, Adv. Astron., 2010, 789293
Despali G., Vegetti S., 2017, MNRAS, 469, 1997
Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, MNRAS, 456, 2486
Despali G., Vegetti S., White S. D. M., Giocoli C., van den Bosch F. C., 2018, MNRAS, 475, 5424
Dolgov A., Hansen S., 2002, Astropart. Phys., 16, 339
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, L64
Fadely R., Keeton C. R., 2012, MNRAS, 419, 936
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Garzilli A., Boyarsky A., Ruchayskiy O., 2017, Phys. Lett. B, 773, 258
Gilman D., Agnello A., Treu T., Keeton C. R., Nierenberg A. M., 2017, MNRAS, 467, 3970
Gilman D., Birrer S., Treu T., Keeton C. R., 2018, MNRAS, 481, 819
Gu L., Kaastra J., Raassen A. J. J., Mullen P. D., Cumbee R. S., Lyons D., Stancil P. C., 2015, A&A, 584, L11
Hezaveh Y. D. et al., 2016, ApJ, 823, 37
Horiuchi S., Humphrey P. J., Oñorbe J., Abazajian K. N., Kaplinghat M., Garrison-Kimmel S., 2014, Phys. Rev. D, 89, 025017
Hsueh J.-W., Fassnacht C. D., Vegetti S., McKean J. P., Spingola C., Auger M. W., Koopmans L. V. E., Lagattuta D. J., 2016, MNRAS, 463, L51
Hsueh J.-W., Despali G., Vegetti S., Xu D., Fassnacht C. D., Metcalf R. B., 2018, MNRAS, 475, 2438
Hsueh J.-W. et al., 2017b, MNRAS, 469, 3713
Iakubovskyi D. A., 2014, Adv. Astron. Space Phys., 4, 9
Iršič V. et al., 2017a, Phys. Rev. D, 96, 023522
Iršič V., Viel M., Haehnelt M. G., Bolton J. S., Becker G. D., 2017b, Phys. Rev. Lett., 119, 031302
Jeltema T., Profumo S., 2016, MNRAS, 458, 3592
Kennedy R., Frenk C., Cole S., Benson A., 2014, MNRAS, 442, 2487
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Koopmans L. V. E., 2005, MNRAS, 363, 1136
Kuzio de Naray R., McGaugh S. S., de Blok W. J. G., 2008, ApJ, 676, 920
Laine M., Shaposhnikov M., 2008, J. Cosmol. Astropart. Phys., 6, 31
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, MNRAS, 439, 300
Lovell M. R. et al., 2016, MNRAS, 461, 60
Lovell M. R. et al., 2017, MNRAS, 468, 4285
Ludlow A. D., et al., 2016, MNRAS, 460, 1214
Mao S., Schneider P., 1998, MNRAS, 295, 587
Marshall H. L., Tananbaum H., Avni Y., Zamorani G., 1983, ApJ, 269, 35
McKean J. P. et al., 2007, MNRAS, 378, 109
Metcalf R. B., 2005, ApJ, 629, 673
Moore B., 1994, Nature, 370, 629
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493 (NFW)
Neronov A., Malyshev D., Eckert D., 2016, Phys. Rev. D, 94, 123504
Nierenberg A. M. et al., 2017, MNRAS, 471, 2224
Polisensky E., Ricotti M., 2011, Phys. Rev. D, 83, 043506
Ringwald A., 2016, in Proceedings of the Neutrino Oscillation Workshop (NOW2016), SISSA Medialab srl, Trieste, Italy. p. 81
Robles V. H. et al., 2017, MNRAS, 472, 2945
Ruchayskiy O. et al., 2016, MNRAS, 460, 1390
Schneider A., 2016, J. Cosmol. Astropart. Phys., 4, 059
Schneider A., Smith R. E., Macciò A. V., Moore B., 2012, MNRAS, 424, 684
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Shi X., Fuller G. M., 1999, Phys. Rev. Lett., 82, 2832
Vegetti S., Koopmans L. V. E., 2009, MNRAS, 392, 945

Vegetti S., Czoske O., Koopmans L. V. E., 2010a, MNRAS, 407, 225
Vegetti S., Koopmans L. V. E., Bolton A., Treu T., Gavazzi R., 2010b, MNRAS, 408, 1969
Vegetti S., Lagattuta D. J., McKean J. P., Auger M. W., Fassnacht C. D., Koopmans L. V. E., 2012, Nature, 481, 341
Vegetti S., Koopmans L. V. E., Auger M. W., Treu T., Bolton A. S., 2014, MNRAS, 442, 2017
Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534
Vogelsberger M., Zavala J., Cyr-Racine F. Y., Pfrommer C., Bringmann T., Sigurdson K., 2016, MNRAS, 460, 1399
Walker M. G., Peñarrubia J., 2011, ApJ, 742, 20
Watson C. R., Li Z., Polley N. K., 2012, J. Cosmol. Astropart. Phys., 3, 018
Xu D. D. et al., 2009, MNRAS, 398, 1235
Xu D., Sluse D., Gao L., Wang J., Frenk C., Mao S., Schneider P., Springel V., 2015, MNRAS, 447, 3189

APPENDIX: THE LIKELIHOOD FUNCTION

To quantify the number of detections and non-detections of perturbers we consider a total number of I bins in mass and J bins in position. The i-th bin in mass corresponds to masses in the range $[m_i^{\rm o}, m_i^{\rm o} + \Delta m^{\rm o}]$, while the j-th bin in position corresponds to projected positions in the range $[x_j^{\rm o}, x_j^{\rm o} + \Delta x^{\rm o}]$.[1] These bins also quantify the non-detections as the number density of perturbers in each bin n_{ji} can be equal to zero.

Here, we derive the Likelihood of detecting $\{n_{ij}\}_{i=1\ \ j=1}^{I\quad J}$ objects. As we are interested in including the contribution from both subhaloes and line-of-sight haloes, n_{ij} is given by the sum of the two populations in each bin ij, i.e. $n_{ij} = n_{s,ij} + n_{l,ij}$. We assume $n_{s,ij}$ and $n_{l,ij}$ to be distributed according to a Poisson probability density with a mean $\mu_{s,ij}$ and $\mu_{l,ij}$, respectively. Taking advantage of the fact that the sum of two Poisson distributed numbers is also Poisson distributed with a mean given by the sum of the two individual means we can write

$$P\left(n_{ij}|\mu_{ij}\right) = \frac{e^{-\mu_{ij}}\mu_{ij}^{n_{ij}}}{n_{ij}!}, \tag{A1}$$

with

$$\mu_{ij} = \mu_{s,ij} + \mu_{l,ij}. \tag{A2}$$

For a given set of the mass function parameters θ, different n_{ij} are conditionally independent of each other, so that

$$\begin{aligned}
&\log\left(P\left(\{n_{ij}\}_{i=1\ \ j=1}^{I\quad J}|\theta\right)\right)\\
&\quad= \log\left(\prod_{i=1}^{I}\prod_{j=1}^{J} P\left(n_{ij}|\theta\right)\right) = \sum_{i=1}^{I}\sum_{j=1}^{J}\log\left(P\left(n_{ij}|\theta\right)\right)\\
&\quad= -\sum_{i=1}^{I}\sum_{j=1}^{J}\mu_{ij} + \sum_{i=1}^{I}\sum_{j=1}^{J}\log\left(\frac{\mu_{ij}^{n_{ij}}}{n_{ij}!}\right).
\end{aligned} \tag{A3}$$

Under the assumptions that the bin widths are small, we can rewrite the first term in the above equation as an integral of the predicted mean number density of subhaloes and line-of-sight haloes (see Section 6.2) over the masses and positions,

$$\begin{aligned}
\sum_{i=1}^{I}\sum_{j=1}^{J}\mu_{ij} &= \sum_{i=1}^{I}\sum_{j=1}^{J}\left[\mu_{s,ij} + \mu_{l,ij}\right]\\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\left[\mu_s\left(m_i^{\rm o}, x_j^{\rm o}\right) + \mu_l\left(m_i^{\rm o}, x_j^{\rm o}\right)\right]{\rm d}m^{\rm o}{\rm d}x^{\rm o}\\
&= \int {\rm d}m^{\rm o}\int {\rm d}x^{\rm o}\left[\mu_s\left(m^{\rm o}, x^{\rm o}\right) + \mu_l\left(m^{\rm o}, x^{\rm o}\right)\right].
\end{aligned} \tag{A4}$$

In particular, given that the number of detected objects is finite we can choose the widths of the bins small enough that the maximum number of detections per bin ij is 1. This assumption is possible because the probability of multiple perturbers matching perfectly in mass and position is insignificantly low. We can then rewrite the second term in equation (A3) as the sum over those bins ij in which a detection with $m_i^{\rm o} = m_k^{\rm ob}$ and $x_j^{\rm o} = x_k^{\rm ob}$ has been made. The detection of n objects then leads to

$$\sum_{k=1}^{n}\log\left[\mu_s\left(m_k^{\rm ob}, x_k^{\rm ob}\right){\rm d}m^{\rm o}{\rm d}x^{\rm o} + \mu_l\left(m_k^{\rm ob}, x_k^{\rm ob}\right){\rm d}m^{\rm o}{\rm d}x^{\rm o}\right], \tag{A5}$$

where all other terms have vanished, since $\log(\frac{\mu_{ij}^0}{0!}) = 0$.

Combining the above considerations, we can express the Likelihood function as

$$\begin{aligned}
&\log\left(P\left(\{n_{ij}\}_{i=1\ \ j=1}^{I\quad J}|\theta\right)\right)\\
&\quad= -\int {\rm d}m^{\rm o}\int {\rm d}x^{\rm o}\left[\mu_s\left(m^{\rm o}, x^{\rm o}\right) + \mu_l\left(m^{\rm o}, x^{\rm o}\right)\right]\\
&\qquad+ \sum_{k=1}^{n}\log\left[\mu_s\left(m_k^{\rm ob}, x_k^{\rm ob}\right){\rm d}m^{\rm o}{\rm d}x^{\rm o} + \mu_l\left(m_k^{\rm ob}, x_k^{\rm ob}\right){\rm d}m^{\rm o}{\rm d}x^{\rm o}\right].
\end{aligned} \tag{A6}$$

For a given set of detections with $m_i^{\rm o} = m_k^{\rm ob}$ and $x_j^{\rm o} = x_k^{\rm ob}$ the number of objects in each bin n_{ij} is determined, hence we can write $P\left(\{n_{ij}\}_{i=1\ \ j=1}^{I\quad J}|\theta\right)$ as a $P\left(\left\{m_1^{\rm ob}, ..., m_n^{\rm ob}\right\}, \left\{\mathbf{x}_1^{\rm ob}, ..., \mathbf{x}_n^{\rm ob}\right\}|\theta\right)$.

[1] For simplicity we have reduced the number of position dimensions to just one. None the less the derivation can be analogously done in two dimensions.

This paper has been typeset from a TEX/LATEX file prepared by the author.